NBER WORKING PAPER SERIES

THE ECONOMIC THEORY OF ILLEGAL GOODS: THE CASE OF DRUGS

Gary S. Becker
Kevin M. Murphy
Michael Grossman

Working Paper 10976
http://www.nber.org/papers/w10976

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2004

The Economic Theory of Illegal Goods: the Case of Drugs
Gary S. Becker, Kevin M. Murphy, and Michael Grossman
NBER Working Paper No. 10976
December 2004
JEL No. D00, D11, D60, I11, I18

ABSTRACT

This paper concentrates on both the positive and normative effects of punishments that enforce laws to make production and consumption of particular goods illegal, with illegal drugs as the main example. Optimal public expenditures on apprehension and conviction of illegal suppliers obviously depend on the extent of the difference between the social and private value of consumption of illegal goods, but they also depend crucially on the elasticity of demand for these goods. In particular, when demand is inelastic, it does not pay to enforce any prohibition unless the social value is negative and not merely less than the private value. We also compare outputs and prices when a good is legal and taxed with outputs and prices when the good is illegal. We show that a monetary tax on a legal good could cause a greater reduction in output and increase in price than would optimal enforcement, even recognizing that producers may want to go underground to try to avoid a monetary tax. This means that fighting a war on drugs by legalizing drug use and taxing consumption may be more effective than continuing to prohibit the legal use of drugs.

Gary S. Becker
Department of Economics
University of Chicago
1126 East 59th Street
Chicago, IL 60637
gbecker@uchicago.edu

Michael Grossman
NBER
365 Fifth Avenue
New York, NY 10016
and CUNY Graduate Center
mgrossman@gc.cuny.edu

Kevin M. Murphy
Graduate School of Business
University of Chicago
Chicago, IL 60637
and NBER
kevin.murphy@gsb.uchicago.edu

1. Introduction

The effects of excise taxes on prices and outputs have been extensively studied. An equally large literature discusses the normative effects of these taxes measured by their effects on consumer and producer surplus. However, the emphasis has been on monetary excise taxes, while non-monetary taxes in the form of criminal and other punishments for illegal production of different goods have been discussed only a little (important exceptions are MacCoun and Reuter, 2001 and Miron, 2001).

This paper concentrates on both the positive and normative effects of punishments that enforce laws to make production and consumption of particular goods illegal. We use the supply and demand for illegal drugs as our main example, a topic of considerable interest in its own right, although our general analysis applies to the underground economy, prostitution, restrictions on sales of various goods to minors, and other illegal activities.

Drugs are a particularly timely example not only because they attract lots of attention, but also because every U.S. president since Richard Nixon has fought this war with police, the FBI, the CIA, the military, a federal agency (the DEA), and military and police forces of other nations. Despite the wide scope of these efforts–and major additional efforts in other nations–no president or drug "czar" has claimed victory, nor is a victory in sight.

Why has the War on Drugs been so difficult to win? How can international drug traffickers command the resources to corrupt some governments, and thwart the extensive efforts of the most powerful nation? Why do efforts to reduce the supply of drugs lead to violence and greater influence for street

gangs and drug cartels? To some extent, the answer lies in the basic theory of enforcement developed in this paper.

Section 2 sets out a simple graphical analysis that shows how the elasticity of demand for an illegal good is crucial to understanding the effects of punishment to producers on the overall cost of supplying and consuming that good. Section 3 formalizes that analysis, and adds expenditures by illegal suppliers to avoid detection and punishment.

That section also derives the optimal public expenditures on apprehension and conviction of illegal suppliers. The government is assumed to maximize a welfare function that takes account of differences between the social and private values of consumption of illegal goods. Optimal expenditures obviously depend on the extent of this difference, but they also depend crucially on the elasticity of demand for these goods. In particular, when demand is inelastic, it does not pay to enforce any prohibition unless the social value is negative and not merely less than the private value.

Section 4 compares outputs and prices when a good is legal and taxed with outputs and prices when the good is illegal. It shows that a monetary tax on a legal good could cause a greater reduction in output and increase in price than would optimal enforcement, even recognizing that producers may want to go underground to try to avoid a monetary tax. Indeed, the optimal monetary tax that maximizes social welfare tends to exceed the optimal non-monetary tax. This means, in particular, that fighting a war on drugs by legalizing drug use and taxing consumption may be more effective than continuing to prohibit the legal use of drugs.

Section 5 generalizes the analysis in sections 2-4 to allow producers to be heterogeneous with different cost functions. Since enforcement is costly, it is efficient to direct greater enforcement efforts toward marginal producers than toward infra-marginal producers. That implies greater enforcement against weak and small producers because marginal producers tend to be smaller and economically weaker. By contrast, if the purpose of a monetary tax partly is to raise revenue for the government, higher monetary taxes should be placed on infra-marginal producers because these taxes raise revenue without much affecting outputs and prices.

Many drugs are addictive and their consumption is greatly affected by peer pressure. Section 6 incorporates a few analytical implications of the economic theory of addiction and peer pressure. They help explain why demand elasticities for some drugs may be relatively high, and why even altruistic parents often oppose their children's desire to use drugs.

Section 7 considers when governments should to try to discourage consumption of goods through advertising, like the "just say no" campaign against drug use. Our analysis implies that advertising campaigns can be useful against illegal goods that involve enforcement expenditures to discourage production. However, they are generally not desirable against legal goods when consumption is discouraged through optimal monetary taxes.

Even though our analysis implies that monetary taxes on legal goods can be quite effective, drugs and many other goods are illegal. Section 8 argues that the explanation is related to the greater political clout of the middle classes.

2. A Graphical Analysis

We first analyze the effects of enforcement expenditures with a simple model of the market for illegal drugs. The demand for drugs is assumed to depend on the market price of drugs that is affected by the costs imposed on traffickers through enforcement and punishment, such as confiscation of drugs and imprisonment. The demand for drugs also depends on the costs imposed by the government on users.

Assume that drugs are supplied by a competitive drug industry with constant unit costs c(E) that depend on the resources, E, that governments devote to catching smugglers and drug suppliers. In such a competitive market, the transaction price of drugs will equal unit costs, or c(E), and the full price of drugs P_e, to consumers will equal c(E) + T, where T measures the costs imposed on users through reduced convenience and/or criminal punishments. Without a war on drugs, T=0 and E=0, so that $P_e = c(0)$. This free market equilibrium is illustrated in Figure 1 at point f.

With a war on drugs focused on interdiction and the prosecution of drug traffickers, E>0 but T=0. These efforts would raise the street price of drugs and reduce consumption from its free market level at f to the "war"

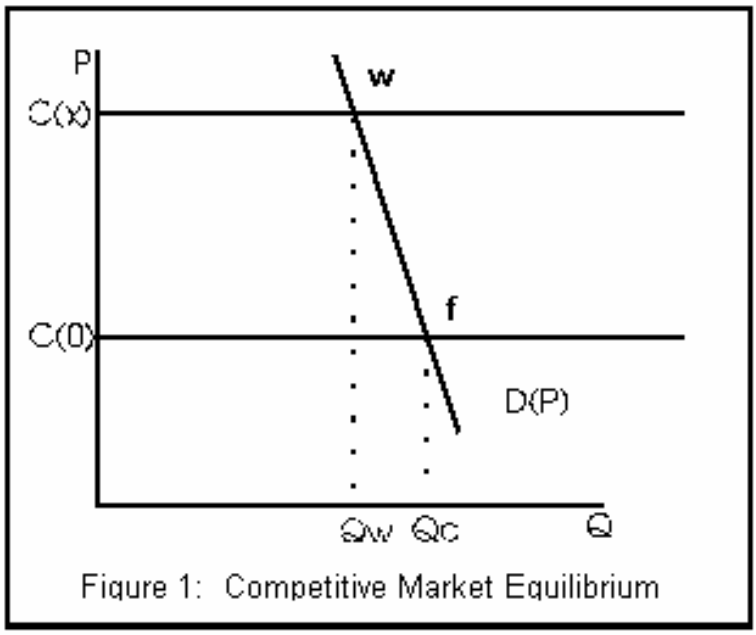

Figure 1: Competitive Market Equilibrium

equilibrium at w, as shown in Figure 1.

This figure shows that interdiction and prosecution efforts reduce consumption. In particular, if Δ measures percentage changes, the increase in costs is given by Δc, and $\Delta Q = \varepsilon \Delta c$, where $\varepsilon < 0$ is the price elasticity of demand for drugs. The change in expenditures on drugs from making drugs illegal is:

$$\Delta R = (1+\varepsilon) \Delta c.$$

When drugs are supplied in a perfectly competitive market with constant unit costs, drug suppliers earn zero profits. Therefore, resources devoted to drug production, smuggling, and distribution will equal the revenues from drug sales in both the free and illegal equilibria. Hence, the change in resources devoted to drug smuggling, including production and distribution,

induced by a "war" on drugs will equal the change in consumer expenditures. Therefore, as eq. (1) shows, total resources devoted to supplying drugs will rise with a war on drugs when demand for drugs is inelastic ($\varepsilon > -1$), and total resources will fall when the demand for drugs is elastic ($\varepsilon < -1$).

When the demand for drugs is elastic, more vigorous efforts to fight the war (i.e. increases in E) will reduce the total resources spent by drug traffickers to bring drugs to market. In contrast, and paradoxically, when demand for drugs is inelastic, total resources spent by drug traffickers will increase as the war increases in severity, and consumption falls. With inelastic demand, resources are actually drawn into the drug business as enforcement reduces drug consumption.

3. The Elasticity of Demand and Optimal Enforcement

This section shows how the elasticity of demand determines optimal enforcement to reduce the consumption of specified goods -again we use the example of illegal drugs. We assume that governments maximize social welfare that depends on the social rather than consumer evaluation of the utility from consuming these goods. Producers and distributors take privately optimal actions to avoid governmental enforcement efforts. In determining optimal enforcement expenditures, the government takes into account how avoidance activities respond to changes in enforcement expenditures.

We use the following notation throughout this section:

Q = consumption of drugs

P = price of drugs to consumers

Demand: $Q = D(P)$

F = monetary equivalent of punishment to convicted drug traffickers

Production is assumed to be CRS. This is why we measure all cost variables per unit output.

c = competitive cost of drugs without tax or enforcement, so $c=c(0)$ from above

A = private expenditures on avoidance of enforcement per unit output

E = level of government enforcement per unit output

$p(E,A)$ = probability that a drug trafficker is caught smuggling, with $\partial p/\partial E > 0$, and $\partial p/\partial A < 0$.

We assume that when smugglers are caught their drugs are confiscated and they are penalized F (per unit of drugs smuggled). With competition and CRS, price will be determined by minimum unit cost. For given levels of E and A, expected unit costs are given by

(2) Expected unit cost $\equiv u = (c + A + p(E,A)\, F) / (1-p(E,A))$.

Working with the odds ratio of being caught rather than the probability greatly simplifies the analysis. In particular, $\theta(E,A) = p(E,A)/(1-p(E,A))$ is this odds ratio, so

(3) $u = (c + A)\,(1+\theta) + \theta\, F$.

Expected unit costs are linear in the odds ratio, θ, since it gives the probability of being caught per unit of drugs sold. Expected unit costs are also linear in the penalty for being caught, F.

The competitive price will be equal to the minimum level of unit cost, or

$$(4a) \quad P = \min_{A} (c + A)(1+\theta) + \theta F.$$

The FOC for cost minimization (with respect to A), taking E and F as given, is

$$(5) \quad -\partial\theta/\partial A \,(c + A + F) = (1 + \theta).$$

We interpret expenditures on avoidance, A, as including the entire increase in direct costs from operating an illegal enterprise. This would include costs from not being able to use the court system to enforce contracts, and costs associated with using less efficient methods of production, transportation, and distribution that have the advantage of being less easily monitored by the government. The competitive price will exceed the costs under a legal environment due to these avoidance costs, A, the loss of drugs due to confiscation, and penalties imposed on those caught.

Hence, the competitive price will equal the minimum expected unit costs, given from eq. (4a) as

$$(4b) \quad P^*(E) = (c + A^*)(1+\theta(E, A^*)) + \theta(E, A^*) F,$$

where A* is the cost minimizing level of expenditures. The competitive equilibrium price, given by this equation, exceeds the competitive equilibrium legal price, c, by A (the added cost of underground production); $(c+A)\theta$, the expected value of the drugs confiscated; and θF, the expected costs of punishment.

An increase in punishment to drug offenders, F, raises the cost and lowers the profits of an individual drug producer. The second order condition for A* in eq. (5) to be a maximum implies that avoidance expenditures increase as F increases. But in competitive equilibrium, a higher F has no effect on expected profits because market price rises by the increase in expected costs due to the higher punishment. In fact, those drug producers and smugglers who manage to avoid apprehension make greater realized profits when punishment increases because the increase in market price exceeds the increase in their unit avoidance costs.

The greater profits of producers who avoid punishment, and even the absence of any effect on expected profits of all producers, does not mean that greater punishment has no desired effects. For the higher market price, given by eq. (4), induced by the increase in punishment reduces the use of drugs. The magnitude of this effect on consumption depends on the elasticity of demand: the more inelastic is demand, the smaller is this effect.

The role of the elasticity and the effect on consumption is seen explicitly by calculating the effect of greater enforcement expenditures on the equilibrium price. In particular, by the envelope theorem, we have[1]

(6a) $dP/dE = \partial\theta/\partial E\ (c + A^* + F) > 0$, and hence

(6b) $d\ln P/d\ln E = \varepsilon_\theta\ \theta\ (c + A^* + F)/P = \varepsilon_\theta\ [\theta(c+A^*+F)/P] = \varepsilon_\theta\ \lambda$

Here, $\lambda = \theta(c+ A^*+ F)/P < 1$, and ε_θ is the elasticity of the odds ratio, θ, with respect to E. Again denoting the elasticity of demand for drugs by ε_d, eq. (6b) implies that

(7) $d\ln Q/d\ln E = \varepsilon_d\ d\ln P/d\ln E = \varepsilon_d\ \varepsilon_\theta\ \lambda < 0.$

If enforcement is a pure public good, then the costs of enforcement to the government will be independent of the level of drug activity (i.e. C(E,Q) =C(E)). On the other hand, if enforcement is a purely private good (with respect to drugs smuggled), an assumption of CRS in production implies that C(E,Q) = QC(E). We adopt a mixture of these two formulations. In addition to these costs, the government has additional costs from punishing those caught. We assume that punishment costs are linear in the number caught

---

[1] Differentiate eq. 4a) with respect to E and note that in general the optimal value of A will vary as E varies:

$$\frac{dP}{dE} = (c + A^* + F)\frac{d\theta}{dE} + \left[(1+\theta) + (c + A^* + F)\frac{d\theta}{dA} \right]\frac{dA}{dE}.$$

From the first order condition for A, the sum of the terms inside the brackets on the right hand side of the equation for dP/dE is zero.

and punished (θQ). With a linear combination of all the enforcement cost components,

(8) $C(Q,E,\theta) = C_1E + C_2QE + C_3\theta Q.$

Eq. (8) implies that enforcement costs are linear in the level of enforcement activities, although they could be convex in E without changing the basic results. Enforcement costs also depend on the level of drug activity (Q), and the fraction of drug smugglers punished (through θ).

The equilibrium level of enforcement depends on the government's objective. We assume that the government wants to reduce the consumption of goods like drugs relative to what they would be in a competitive market. We do not model the source of these preferences, but assume a "social planner" who may value drug consumption by less than the private willingness to pay of drug users, measured by the price, P. If V(Q) is the social value function, then $\partial V/\partial Q \equiv V_q \leq P$, with V_q strictly $< P$ if there is a perceived externality from drug consumption, and hence drug consumption is socially valued at strictly less than the private willingness to pay. When $V_q < 0$, the negative externality from consumption exceeds the positive utility to consumers.

With these preferences, the government chooses E to maximize the value of consumption minus the sum of production and enforcement costs. Thus it chooses E to solve

(9) $\max_{E} W = V(Q(E)) - u(E)Q(E) - C(Q(E), E, \theta(E, A^*(E)))$.

The government incorporates into its decision the privately optimal change in avoidance costs by drug producers and smugglers to any increase in enforcement costs. With the assumption of CRS on the production side, then $u(E)Q(E) = P(E)Q(E)$, and we assume C is given by eq. (8). Thus the planner's problem simplifies to

(10) $\max_{E} W = V(Q(E)) - P(E)\,Q(E) - C_1E - C_2Q(E)E - C_3\theta(E, A^*(E))Q(E)$

The first order condition is

(11) $V_q\, dQ/dE - MR\, dQ/dE - C_1 - C_2\,(Q + (dQ/dE)E) - C_3\left[\theta\frac{dQ}{dE} + Q\left(\frac{\partial\theta}{\partial E} + \frac{\partial\theta}{\partial A}\bullet\frac{dA}{dE}\right)\right] = 0 \rightarrow$

(12a) $C_1 + C_2\,(Q + EdQ/dE) + C_3\,(\theta\, dQ/dE + Q\, d\theta/dE) = V_q\, dQ/dE - MR\, dQ/dE$,

where $MR \equiv d(PQ)/dQ$ denotes marginal revenue.

The left hand side of eq. (12a) is the marginal cost of enforcement, including the effects on output and the odds ratio. The right hand side is the marginal benefit of the reduction in consumption, including the effect on production costs. This equation becomes more revealing if we temporarily assume that

marginal enforcement costs are zero. Then the RHS of this equation equals zero, which simplifies to

12b) $V_q = MR \equiv P(1+1/\varepsilon_d)$, or $V_q/P = 1+1/\varepsilon_d$,

and V_q/P is the ratio of the social marginal willingness to pay to the private marginal willingness to pay of drug users (measured by price).

If $V_q \geq 0$, so that drug consumption has non-negative marginal social value, and if demand is inelastic, so that $MR < 0$, eq. (12b) implies that optimal enforcement would be zero, and free market consumption would be the social equilibrium. There is a loss in social utility from reduced consumption since the social value of additional consumption is positive - even if it is less than the private value–while production and distribution costs increase as output falls when demand is inelastic.

The conclusion that with positive marginal social willingness to pay–no matter how small–inelastic demand, and punishment to traffickers, the optimal social decision would be to leave the free market output unchanged does not assume the government is inefficient, or that enforcement of these taxes is costly. Indeed, the conclusion holds in the case we just discussed where governments are assumed to catch violators easily and with no cost to themselves, but costs to traffickers. Costs imposed on suppliers bring about the higher price required to reduce consumption. But since marginal revenue is negative when demand is inelastic, total costs would rise along with revenue as price rises and output is reduced, while total social value would

fall as output falls if V_q were positive. The optimal social decision is clearly then to do nothing, even if consumption imposes significant external costs on others.

This result differs radically from well-known optimal taxation results with monetary taxes. Then, if the monetary tax is costless to implement, and if the marginal social value of consumption is less than price–no matter how small the difference–it is always optimal to reduce output below its free market level.

Even if demand is elastic, it may not be socially optimal to reduced output if consumption of the good has positive marginal social value. For example, if the elasticity is as high as –11/2, eq. (12b) shows that it is still optimal to do nothing as long as the ratio of the marginal social to the marginal private value of additional consumption exceeds 1/3. It takes very low social values of consumption, or very high demand elasticities, to justify intervention, even with negligible enforcement costs.

Intervention is more likely to be justified when $V_q < 0$: when the negative external effects of consumption exceed the private willingness to pay. If demand is inelastic, marginal revenue is also negative, and eq. (12b) shows that a necessary condition to intervene in this market is that marginal social value be less than marginal revenue at the free market output level.
There are no reliable estimates of the price elasticity of demand for illegal drugs, mainly because data on prices and quantities consumed of illegal goods are scarce. However, estimates generally indicate an elasticity of less than one in absolute value, although one or two studies estimate a larger

elasticity (see Caulkins, 1995, van Ours, 1995). Moreover, few studies of drugs have utilized the theory of rational addiction, which implies that long run elasticities exceed short run elasticities for addictive goods (see section 6).[2] Since considerable resources are spent fighting the war on drugs and reducing consumption, the drug war can only be considered socially optimal with a long run demand elasticity of about –1/2 if the negative social externality of drug use is more than twice the positive value to drug users. Of course, perhaps the true elasticity is much higher, or the war may be based on interest group power rather than maximizing social welfare (see section 8).

Punishment to reduce consumption is easier to justify when demand is elastic and hence marginal revenue is positive. If enforcement costs continue to be ignored, total costs of production and distribution must then fall as output is reduced. If $V_q < 0$, social welfare would be maximized by eliminating consumption of that good because costs decline and social value rises as output falls. However, even with elastic demand and negative marginal social value, rising enforcement costs as output falls could lead to an internal equilibrium.

Figure 2 illustrates another case where it may be optimal to eliminate consumption (ignoring enforcement costs). In this case, demand is assumed

[2] Grossman and Chaloupka (1998) present a variety of estimates of rational addiction models of the demand for cocaine by young adults in panel data. They emphasize an estimate of the long-run price elasticity of total consumption (participation multiplied by frequency given participation) of -1.35. When, however, they include individual fixed effects to control for unmeasured area-specific effects that may be correlated with price and consumption, the elasticity becomes -0.67. One problem with the latter estimate is that biases due to random measurement error in the price of cocaine are exacerbated in the fixed-effects specification.

to be elastic, and at the free market equilibrium, V_q is positive and greater than MR, but it is less than the free market price. MR is assumed to rise more rapidly than V_q does as output falls, so that they intersect at Q_u. That point would equate MR and V_q, but it violates the SOC for a social maximum.

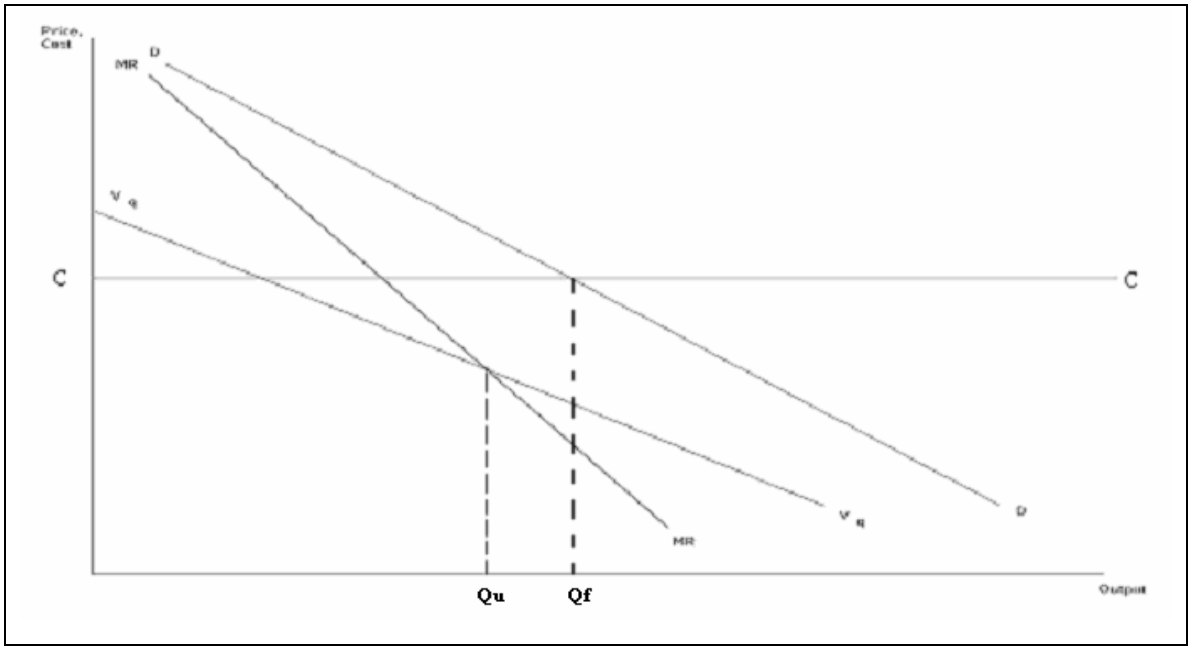

Figure 2

The optimum in this case is to go to one of the corners, and either do nothing and remain with the free market output, or fight the war hard enough to eliminate consumption. Which of these extremes is better depends on a comparison of the area between V_q and MR to the left of Q_u, with the corresponding area to the right. If the latter is bigger, output remains at the free market level, even if the social value of consumption at that point were much less than its private value. It would be optimal to remain at the free

market output when reducing output from the free market level lowers social value by sufficiently more than it lowers production costs.

Eq. (12a) incorporates enforcement costs into the first order conditions for a social maximum. It is interesting that marginal enforcement costs also depend on the elasticity of demand, and they too are greater when demand is more inelastic. To see this, rewrite the LHS of eq. (12a) as

$$MC_E = C_1 + C_2Q + C_2E dQ/dE + C_3(\theta\, dQ/dE + Q d\theta/dE)$$

$$= C_1 + C_2Q\,(1 + dlnQ/dlnE) + C_3(\theta dQ/dE + Q d\theta/dE)$$

$$= C_1 + C_2Q\,(1 + dlnQ/dlnE) + C_3\theta\, Q/E(dlnQ/dlnE + \varepsilon_{\theta *})$$

$$= C_1 + C_2Q\,(1 + \lambda\, \varepsilon_\theta\, \varepsilon_d) + C_3\theta Q/E\, \varepsilon_{\theta *}\,(1 + \lambda\, \varepsilon_d\, \varepsilon_\theta/\varepsilon_{\theta *}). \quad (13)$$

Here $\varepsilon_{\theta *}$ is the total elasticity of θ with respect to E, which includes the indirect effect of E on the privately optimal changes in avoidance costs, A, by producers and distributors. That is, since

$$d\theta/dE = \partial\theta/\partial E + (\partial\theta/\partial A)(dA/dE) \rightarrow \varepsilon_{\theta *} = \varepsilon_\theta + \varepsilon_A\, dlnA/dlnE.$$

Eq. (13) shows that marginal enforcement costs are greater, the smaller is ε_d in absolute value because consumption falls more rapidly as enforcement increases when demand is more elastic. Since expenditures on apprehension and punishment depend on output, a slower fall in output with more inelastic

demand causes enforcement expenditures to grow more rapidly. Indeed, eq.(13) implies that if demand is sufficiently elastic, marginal enforcement costs can be negative when enforcement increases since the drop in the scale of production can more than offset the increased cost per unit.

So the elasticity of demand is key on both the cost and benefit sides of enforcement. When demand is elastic, total industry costs fall as consumption is reduced, and enforcement costs increase more slowly, or they may even fall. Extensive government intervention in this market to reduce output would then be attractive if the marginal social value of consumption is low. In contrast, when demand is inelastic, total production costs rise as consumption falls, and enforcement costs rise more rapidly. With inelastic demand, a war to reduce consumption would be justified only when marginal social value is very negative. Even then, such a war will absorb a lot of resources.

4. A Comparison with Monetary Taxes

It is instructive to compare these results for enforcement effects with well-known results for monetary taxes on legal goods. The social welfare function for these monetary taxes that corresponds to the welfare function for enforcement of the prohibition against drugs in eq. 9 is, ignoring avoidance and enforcement costs,

(14a) $W_m = V(Q) - cQ - (1- \delta)\tau Q,$

where τ is the monetary tax per unit output of drugs, and δ gives the value to society per each dollar taxed away from taxpayers. Since in competitive equilibrium $P = c + \tau$, eq. (14a) can be rewritten as

(14b) $W_m = V(Q) - cQ - (1 - \delta)(P(Q)Q - cQ)$

The first order condition for Q is

(15a) $V_q = c + (1 - \delta)(MR - c)$,

or

(15b) $\tau = P - V_q + (1-\delta)\left(P(1+\frac{1}{\varepsilon_d}) - c \right)$

If tax receipts are a pure transfer, so that $\delta=1$, eq. (15a) or (15b) gives the classical result that the optimal monetary tax equals the difference between marginal private (measured by P) and marginal social value. With a pure transfer, the elasticity of demand is irrelevant. The optimal monetary tax is positive if the marginal social value of consumption at the free market competitive position is less than the competitive price.

The elasticity of demand becomes relevant if there are net social costs or benefits from the transfer of resources to the government. If government tax receipts are socially valued at less than dollar for dollar ($\delta<1$), and if demand is inelastic ($\varepsilon_d > -1$), the optimal tax would be positive only if the marginal

social value of consumption were sufficiently less than the marginal private value. The converse holds if tax revenue is highly valued so that $\delta > 1$. The optimal tax on this good might then be positive, even if demand is inelastic and social value exceeds private value.

Of course, if the monetary tax gets too high, some drug producers might try to avoid the tax by trafficking in the underground economy. An optimal monetary tax on a legal good is still always better than optimal enforcement against an illegal good. The proof assumes that the government can choose optimal punishments for producers who sell in the underground economy, and that demand for the good is not reduced by making the good illegal. Let E^* denote the optimal value of enforcement that maximizes the government's welfare function given by eq. (10), and recall that this optimal value takes account of avoidance expenditures by producers. Then, from eq. (4b), the optimal price is $P^* = (c + A^*)(1 + \theta(E^*, A^*)) + \theta(E^*, A^*)F$.

Assume that enforcement against drug producers who try to avoid the monetary tax by selling in the underground economy is sufficient to raise the unit costs of these producers to the same P^*. If the monetary tax is then set at slightly less than $\tau^* = P^* - c$, firms that produce in the legal sector will be slightly more profitable than illegal underground firms. The latter would be driven out of business, or become legal producers. Even ignoring the revenue from the monetary tax, enforcement costs would then be lower with this monetary tax than with optimal enforcement since few would produce illegally. Indeed, in this case, governments only have to incur the fixed

component of enforcement costs, C_1 E*, since in equilibrium no one produces underground.

The government could even enforce an optimal monetary tax that raises market price above the price with optimal enforcement when drugs are illegal. This is sometimes denied with the argument that producers would go underground if monetary taxes are too high. But the logic of the analysis above on deterring underground production shows that this claim is not correct. Whatever the level of the optimal monetary tax, it could be enforced by raising punishment and apprehension sufficiently to make the net price to producers in the illegal sector below the legal price with the optimal monetary tax. Since no one would then produce in the illegal sector, actual enforcement expenditures would be limited to the fixed component, C_1 E*.

To be sure, the optimal monetary tax would depend on this fixed component of enforcement expenditures. But perhaps the most important implication of this analysis relates to a comparison of optimal monetary taxes and enforcement against illegal goods. If enforcement costs are ignored, and if $\delta > 0$, a comparison of the FOC's in eqs. (12b) and (15a) clearly shows that the optimal monetary tax would exceed the optimal "tax" due to enforcement and punishment if demand were inelastic since marginal revenue is then always less than c, unit legal costs of production. The incorporation of enforcement costs only reinforces this conclusion about a higher monetary tax since enforcement costs of cutting illegal output are greater when all production is illegal rather than when some producers go underground to avoid monetary taxes.

If δ=1 and there are no costs of enforcing the optimal monetary tax, optimal output (Q_f) satisfies $V_q = c$ (see eq. (15a)). When some enforcement costs must be incurred to insure that no one produces underground, optimal output (Q^*) satisfies

16) $(V_q - c)dQ/dE = C_1$.

Since an increase in E lowers Q, V_q must be less than c. That implies that Q^* exceeds Q_f. Note that optimal legal output is zero when V_q is negative, and there are no enforcement costs. But eq. (16) could be satisfied at a positive output level when V_q is negative as long as dQ/dE is sufficiently negative at that output.

Various wars on drugs have been only partially effective in cutting drug use, but the social cost has been large in terms of resources spent, corruption of officials, and imprisonment of many producers, distributors, and drug users. Even some individuals who are not libertarians have called for decriminalization and legalization of drugs because they believe the gain from these wars has not been worth these costs. Others prefer less radical solutions, including decriminalization only of milder drugs, such as marijuana, while preserving the war on more powerful and more addictive substances, such as cocaine.

Our analysis shows, moreover, that using a monetary tax to discourage legal drug production could reduce drug consumption by more than even an efficient war on drugs. The market price of legal drugs with a monetary excise tax could be greater than the price induced by an optimal war on

drugs, even when producers could ignore the monetary tax and consider producing in the underground economy. Indeed, the optimal monetary tax would exceed the optimal price due to a war on drugs if the demand for drugs is inelastic- as it appears to be- and if the demand function is unaffected by whether drugs are legal or not- the evidence on this is not clear. With these assumptions, the level of consumption that maximizes social welfare would be smaller if drugs were legalized and taxed optimally instead of the present policy of trying to enforce a ban on drugs.

5. Heterogeneous Taxes and Suppliers

The assumptions made so far of identical firms and of a constant enforcement tax per unit of output has brought out important principles that mainly continue to hold more generally. This section deals briefly with a few novel aspects of optimal enforcement when producers have different costs.

The US experience with the prohibition of alcoholic beverages shows that most companies which produced the good when it was legal exited the industry after prohibition. Legal producers of beer and other alcoholic beverages were replaced by companies who were more willing to, and more skilled at, delivering beer and liquor to underground illegal retailers, while evading or bribing the police and courts that enforced prohibition. More generally, suppliers of illegal goods would generally differ from those who would produce and sell the goods when they were illegal.

Presumably, illegal firms would have higher production costs under the contractual and other aspects of the legal and economic environment when production is legal than the firms that produced the goods when they were

legal. Otherwise, producers under prohibition would have been the low cost producers, and they would have dominated the legal industry.

By limiting the firms that want to enter, prohibition of a good is likely to lower the elasticity of supply. If the elasticity were less than infinite because some firms are relatively low cost producers in an illegal environment, the government should be more active in its enforcement against marginal producers and marginal outputs. Any real expenditures on more efficient infra-marginal producers and infra-marginal units is a waste and serves no efficiency purpose.

With heavier enforcement against marginal producers, the change in producer costs is less than the change in consumer expenditures as the equilibrium price is forced up by enforcement activities. Social costs would then be measured by the smaller rise in producer costs, not by the larger rise in consumer expenditures, as long as the increase in producer rents or profits are considered a transfer from consumers to producers, and not a social cost of the reduction in consumption. However, if no social value were placed on these profits–such as profits to a drug cartel–social cost would still be measured by consumer expenditures, and it would not then be possible to reduce social costs by enforcing more intensively against marginal producers than against more efficient producers.

Of course, to go after marginal producers more heavily requires information on the costs of different producers in an illegal environment. Although the direct information on such costs may be limited, indirect evidence may be considerable since marginal firms tend to be smaller, younger, less

profitable, and financially weaker. It would then be optimal to impose higher unit taxes on smaller, younger, and weaker suppliers. Weaker enforcement against larger producers of drugs is often taken as evidence that these producers bribed and corrupted police and other officials–which may be true. At the same time, our analysis shows that such weaker enforcement may be socially optimal. Government policy should recognize that heavy enforcement against larger and more efficient producers is a wasteful way to raise price and reduce consumption of drugs.

Note the contrast with well-known results on optimal monetary taxation of heterogeneous producers. If tax revenue is highly valued, higher monetary taxes should be extracted from infra-marginal producers than from marginal producers because more efficient producers collect profits that can be taxed away without adverse effects on their incentives. In the extreme case of completely inelastic supply, monetary taxes have no effects on incentives or output, and produce abundant tax revenue.

6. Addictions and Peer Pressure

Some drugs are highly addictive, although the degree of addiction of many of them is controversial. Most drug users start in their teens or early twenties, and peer pressure is especially strong among teenagers (see e.g., Coleman, 1961). This is why it is important to integrate both peer pressure and addiction into an analysis of the positive and normative aspects of illegal markets for drugs.

The combination of addiction to a good and peer pressure to consume that good may lower the short run elasticity of demand for drugs, but they raise

its long-run response to price and other shocks that are common to different consumers. These forces may raise the long run elasticity of demand for drugs to sizeable level, although not necessarily greater than one. For example, essentially all estimates of the long run demand elasticity are less than one for a highly addictive good like smoking, which is apparently also greatly affected by peer pressure.

Some models of addiction imply that individuals consume greater quantities of addictive goods than they "really" would like to. The usual claim is that multiple inconsistent selves battle for control over an individual's decision-making process, such as when they use hyperbolic discounting of future utilities. The implications for optimal excise taxes of these models are generally not unique to harmful addictive goods, and apply to the consumption of all goods that trade off present utility for future disutility. If these approaches are correct, they would provide additional reasons why the utility from the social consumption of harmful addictive goods is below private utilities.

Even if increased consumption of a good by members of a peer group lowers the utility of other members, that could stimulate greater consumption of this good by all other members through raising the good's marginal utility to these members. In this case, goods that are sensitive to peer pressure, such as drugs, would be consumed excessively from the viewpoint of members of the peer group as well. This would be a further reason why the social value of the consumption of drugs was below the private values of individuals. Of course, if greater consumption by peers raised rather than lowered utilities

of other members, social utility would exceed private utilities due to the effects of peer pressure.

If parents believe their children use drugs because of the negative influence of peer pressure, this analysis provides one reason why even altruistic parents try to reduce drug use by their children. Promoting effective reductions in drug use by making drugs illegal or placing a high excise tax on drugs would then raise the utilities of their children and other members of the peer group. Also, altruistic parents may be concerned about their children consuming addictive goods that lower the children's future utilities because altruistic parents may be "forced" to help them out in the future when the children's utilities are lower.

7. Just Say No

Monetary excise tax theory leaves little room for government policies to reduce the demand function for goods that are taxed**.** If the purpose is to raise revenue, why try to reduce demand that reduces tax revenue? In addition, it is more efficient to cut consumption because of an externality with optimal monetary taxes that also raise revenue than with costly programs that reduce the demand function.

These advantages do not apply to illegal goods with enforcement and punishment costs. These expenditures could be reduced by successful government efforts to discourage consumption of these goods. The campaign to "just say no" to drugs is an example of an attempt to reduce consumption.

Illegal goods like drugs have two classes of policy instruments: enforcement and punishment strategies that reduce consumption by raising the real costs and prices of supplying the goods, and expenditures on "education," "advertising," and " persuasion" that reduce demand for these goods. If π represents these expenditures, the social value function W in eq. (10) would be modified to

$$W= V(Q(E, \pi), \pi) - P(E)Q(E, \pi) - c(\pi).$$

In this equation, $c(\pi)$ is the cost of producing π units of persuasion against consuming Q, and for simplicity we ignore enforcement costs (C). We allow W to depend directly on π as well as indirectly through π's effect on Q.

The FOC for maximizing W with respect to π is

$$(18) \quad - Q_\pi(P - V_q) + V_\pi = c_\pi,$$

where a subscript denotes a partial derivative.

The term on the RHS of this equation, $c_\pi > 0$, gives the marginal cost of producing π, and the LHS gives the marginal benefit of additional π. If persuasion is effective in reducing consumption then $Q_\pi < 0$. Reduction in consumption is desirable if the marginal social value of consumption, V_q, is less than its private value, measured by P. The sign of the term V_π is positive or negative as society likes or dislikes the "persuasion". However,

persuasion can have social value even if it is disliked because the LHS of eq. (18) can be positive, even if $V_\pi < 0$, if V_q is sufficiently less than P.

What is interesting about the FOC for persuasive activities to reduce demand is that these activities may be effective in raising social welfare when enforcement activities are least effective. We have shown that it is socially optimal not to spend resources to reduce consumption of an illegal good if its demand is inelastic, and if the marginal social value of its output is positive ($V_q > 0$).

Eq. (18) shows, however, that the elasticity of demand has no effect on the effectiveness of persuasive activities to reduce consumption of an illegal good. Therefore, even if demand is inelastic, and even if the marginal social value of its consumption were positive, there still could be a strong case for persuasive efforts to reduce consumption of an illegal good. This depends on whether $V_q < P$, or whether marginal social value is less than private value. If it is less, persuasion would raise social welfare if it is cheap to produce, and if persuasion efforts do not have a large negative social value. Note that $V_q < P$ is the same criterion that determines whether monetary taxes are desirable.

Persuasion may also raise the effectiveness of enforcement expenditures by raising the elasticity of demand. Becker and Murphy (1992) show that advertising tends to raise the elasticity of demand because it tries to target marginal consumers and increase their demands. It is more efficient for governments to try to reduce the demand for illegal goods of marginal consumers than of other consumers since the former are easier to affect

because they get little surplus from consuming these goods. This means that persuasion does not have to reduce their willingness to pay by a lot to discourage them from consuming these goods. Persuasion could be an effective instrument of government policy not only by reducing demand for illegal goods, but also by raising the effectiveness of enforcement through raising the elasticity of demand for these goods.

8. Why Are Goods Illegal Rather Than Legal And Taxed?

We demonstrated that if the social value of a good is less than its private value, it would be most effective to allow the good to be legal, and impose the right monetary tax to account for the discrepancy between private and social values. Yet throughout history goods like drugs, prostitution, and gambling have frequently been illegal. One answer to this discrepancy between actual and optimal policies depends on their different impacts on the consumption of middle class and poorer persons. Higher and middle level income families often prefer certain goods to be illegal rather than taxed, while poorer persons prefer the opposite. If the poor have much less political power, these goods would end up being illegal.

Even if the increase in money price were the same when a good was illegal and when it was legal and taxed, the consumption of richer and poorer consumers would be affected differently. Suppose a monetary tax raises the price of a good by ΔP to all consumers, and that appropriate enforcement policies prevented a black market in the good. This price increase will tend to have different income and substitution effects to members of different income groups. Even if preferences did not differ by income class, the poor would be more affected by a monetary price increase when the income

elasticity of demand is less than one, and when the value of the time spent consuming the good is a relatively large part of the total cost of consumption. Estimated income elasticities for cocaine, marijuana, and heroin are generally much less than one (see Grossman and Chaloupka, 1998, Liccardo Pacula, Grossman, Chaloupka, O'Malley, Johnston, and Farrelly, 2000, and Saffer and Chaloupka, 1999. However, van Ours, 1995, finds a high income elasticity for opium in the Dutch West Indies).

Up to a point the income and substitution effects work in the same way when the street price of drugs rises because it is illegal. However there is price discrimination when goods are illegal because the total price of illegal goods tends to be lower to poorer persons. Since most crimes are concentrated in poorer neighborhoods, illegal drug production and distribution also tends to be concentrated in these neighborhoods. This makes illegal goods cheaper to persons who live in these neighborhoods since access to them is easier. The total cost of drugs and other illegal goods is cheaper to poorer persons also because they are more likely to be involved in the trafficking in these goods. They are more involved because the cost of imprisonment and similar punishments from selling drugs is less to individuals with lower opportunities in the legal sector. The full cost argument is stronger if we consider enforcement against consumers. Since the non-monetary tax, i.e., punishment, is more time intensive, this corresponds to a difference in the value of the tax between classes that exacerbates the effect. There are also reputational effects that make conviction costlier for the wealthy. In fact, more than half of all persons imprisoned on drug charges are African-American (see Maguire and Pastore, 2001, and Harrison and Beck, 2003).

Even disclosure of use sometimes is very costly to higher income and more educated persons. During his first presidential campaign, Bill Clinton had to deny that he inhaled on the allegedly few occasions when he smoked marijuana. Marijuana use during his student days cost Judge Douglas Ginsberg a Supreme Court seat.

Our conclusion is that making goods illegal and punishing suppliers and consumers by imprisonment and other methods are more costly to higher income persons, and hence tends to reduce their consumption more than consumption of lower income persons. Even if low, middle, and higher income parents have the same desire to discourage drug use by their children, the great political influence of higher education and income groups would explain why drugs are illegal rather than subject to sizeable monetary excise taxes. It also helps explain why punishment is mainly imposed on suppliers rather than consumers of drugs since traffickers are more likely than consumers to be low-income persons.

This analysis also helps explain why prostitution and much gambling are illegal rather than legally consumed with high excise taxes. If individuals at all income levels want to discourage consumption of these goods by children and other family members or friends, the politically powerful middle and higher income persons would prefer to make them illegal rather than legal and subject to high "sin" taxes. The explanation is again that consumption of these goods by middle and richer individuals are reduced more when they are illegal than subject to the high sin taxes. The intent may not be to inflict greater harm on the poor, but making goods like drugs, gambling, and

prostitution illegal, and mainly punishing traffickers, has precisely that effect.

Acknowledgments

Our research has been supported by The Robert Wood Johnson Foundation (Grant I.D. # 045566 to the National Opinion Research Center), the Hoover Institution Project on Drugs, and the Stigler Center for the Study of the Economy and the State. Helpful comments were received from Steve Levitt and Ivan Werning, and at seminars at The University of Chicago and Harvard University. Steve Cicala provided excellent research assistance. This paper has not undergone the review accorded official National Bureau of Economic Research publications; in particular, it has not been submitted for approval by the Board of Directors. Any opinions expressed are ours and not those of the NBER, the Robert Wood Johnson Foundation, NORC, the Hoover Institution, or the Stigler Center.

References:

Becker, Gary S., and Murphy, Kevin M., "A Simple Theory of Advertising as a Good or Bad." Quarterly Journal of Economics, 108, No. 4 (November 1993), 941-964.

Caulkins, Jonathan, P., Estimating the Elasticities of Demand for Cocaine and Heroin with Data from 21 Cities from the Drug Use Forecasting (DUF) Program, 1987-1991 [Computer file]. ICPSR version. Santa Monica, CA: RAND (Corporation [producer], 1995. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 1996.

Coleman, James S., *The Adolescent Society*. Free Press. New York: 1961

Grossman, Michael and Chaloupka, Frank J., "The Demand for Cocaine by Young Adults: A Rational Addiction Approach," Journal of Health Economics, 17, No. 4, August 1998, 427-474).

Harrison, Paige M. and Beck, Allen J., *Prisoners in 2002*. The Bureau of Justice Statistics, Washington, D.C.: 2003, P. 10.

MacCoun, Robert and Reuter, Peter, Drug War Heresies: Learning from Other Vices, Times and Places (Cambridge, U.K.: Cambridge University Press, 2001).

Maguire, Kathleen and Pastore, Ann L., eds. *Sourcebook of Criminal Justice Statistics*. The Bureau of Justice Statistics, Washington, D.C.: 2001, p 515.

Jeffrey A. Miron, "Drug War Crimes", The Independent Institute, 2004.

Pacula, Rosalie Liccardo, Grossman, Michael, Chaloupka, Frank J., O'Malley, Patrick M., Johnston, Lloyd D., and Farrelly, Matthew C. "Marijuana and Youth" in Risky Behavior Among Youths: An Economic Analysis, edited by Jonathan Gruber, University of Chicago Press, 2000, pp. 271-326.

Saffer, Henry and Chaloupka, Frank J. "The Demand for Illicit Drugs," Economic Inquiry, 37, No. 3, July 1999, pp. 401-411.

van Ours, Jan C. “The Price Elasticity of Hard Drugs: The Case of Opium in the Dutch East Indies, 1923-28” Journal of Political Economy, 103, No. 2, April 1995, 261-279.